Szabó, Beáta; Lengyel, Edina; Padisák, Judit; Stenger-Kovács, Csilla. Benthic diatom metacommunity across small freshwater lakes: driving mechanisms, β-diversity and ecological uniqueness. HYDROBIOLOGIA 828: pp. 183-198. (2018)

Benthic diatom metacommunity across small freshwater lakes: driving mechanisms, β-diversity and ecological uniqueness

Beáta Szabó[1,2*], Edina Lengyel[1], Judit Padisák[1,2] Csilla Stenger-Kovács[2]

[1]MTA-PE Limnoecology Research Group, Hungarian Academy of Sciences, Egyetem str. 10, H-8200 Veszprém, Hungary

[2] University of Pannonia, Department of Limnology, Egyetem str. 10, H-8200 Veszprém, Hungary

*corresponding author: e-mail: szabobea@almos.uni-pannon.hu, phone number: +36-88-624337, ORCID: http://orcid.org/0000-0002-2972-0306

Abstract

In this study, driving forces and diversity patterns of a benthic diatom metacommunity across small freshwater lakes exhibiting environmental heterogeneity were investigated. Furthermore, local (LCBD) and species (SCBD) contributions to β-diversity and their driving parameters were assessed with abundance- and incidence-based analyses. Our results revealed that both spatial distance and environmental heterogeneity affected the community assembly, which corresponds most to the mass-effect (ME) concept. This theory was confirmed by high α-diversity of sampling sites, however, high overall β-diversity enhanced mainly by turnover contradicted the ME paradigm. LCBD indices were affected by environmental variables furthermore, LCBD and LCBD in terms of species replacement showed a strong positive correlation. The ecologically most unique sites hosted relatively low species richness, and common species with intermediate-sized or broad niches contributed mostly to the regional β-diversity. However, abundance- and incidence-based calculations revealed different relationships of SCBD with the species' total abundance and the number of occupied sites. Consequently, we favor the previous suggestions that comprehensive research focusing on conservation should incorporate the investigation of LCBD, SCBD, species-rich sites and also ecologically restricted species. Moreover, in assessing ecological uniqueness, both abundance and binary data sets should be considered since they might shed light on distinct patterns.

Key words: assembly mechanisms, diversity patterns, ecological uniqueness, mass-effect, species richness

Introduction

The current ecology- and conservation-oriented research tends to explore the possible causes of community assembly by examining it at regional scale, rather than by only "snap-shot" investigation of groups of biota within a given habitat. That is, studies focusing on metacommunity processes as well as β-diversity and its components are gaining more and more attention. Within the metacommunity framework (Leibold et al., 2004) four different concepts can be distinguished in explaining the importance of local- (species' competitive abilities, demographic processes) and regional-scale (degree of environmental heterogeneity, dispersal) processes. In the neutral theory (NT), species are assumed to be identical concerning their interspecific interactions and response to any limiting factor; demographic processes (birth-death rates) are stochastic; the environment is homogeneous in the region; and species are limited in their dispersion. The patch dynamic (PD) archetype assumes that the species' relative competitive abilities depend on the local environmental conditions; the population-level extinctions are stochastic due to the individual-level stochasticity; the environment is completely homogeneous or spatial heterogeneity may occur in response to the environment; dispersal is limited but interspecific differences in colonization abilities are allowed. In the mass-effect (ME) concept, competitive abilities and birth-death rates are assumed to be largely dependent on the local environment, which displays heterogeneous patterns; species are able to persist in suboptimal localities if there is a sufficient immigration from adjacent sites with high population growth. The species-sorting (SS) concept, similarly to the ME, expects that the environment is heterogeneous, local conditions regulate the competitive abilities of species and the demographic processes; dispersal is sufficient, thus each species can persist in any habitat where it can achieve positive population growth (Leibold & Chase, 2017). Processes assumed to be acting in the four metacommunity archetypes is summarized in Fig. 1. However, the role of these local- and regional-scale

processes, and thus the interpretation of metacommunity concepts, may change with the extent of the investigated area (Langenheder & Ragnarsson, 2007; Mykrä et al., 2007; Heino et al., 2010; Vilmi et al., 2016) and the connectivity among sites (Göthe et al., 2013; Dong et al., 2016; Vilmi et al., 2016).

In estimating the heterogeneity of communities and in unraveling the mechanisms acting behind metacommunity patterns, β-diversity analyses play a key role (Viana et al., 2016). One of the most important and most commonly applied framework for β-diversity surveys was proposed by Baselga (2010). He introduced the multiple-site Sørensen dissimilarity index as suitable to measure overall dissimilarity among a set of sampling sites, which can be divided into turnover (species replacement) and nestedness (reflects species loss) components (Baselga et al., 2007; Baselga, 2010). Its analogous method, the abundance-based multiple-site Bray-Curtis dissimilarity index, has been published recently and can be partitioned into abundance balanced variation and abundance gradients components (Baselga, 2017).

Total β-diversity (i.e. the total variation in community concerning binary or abundance matrix) can be divided into the relative contribution of individual sampling units (Local Contribution to Beta Diversity - LCBD) and of individual species (Species Contribution to Beta Diversity - SCBD) to the overall β-diversity, which targets the assessing of ecological uniqueness of sites and species (Legendre & De Cáceres, 2013). In addition, calculations have been extended to the measure of sites' uniqueness in terms of species replacement and nestedness (Legendre & De Cáceres, 2013).

Although Baselga's (2010) incidence-based calculations are widely used in terrestrial and aquatic ecology (e.g., Maloufi et al., 2016; Conradi et al., 2017; Szabó et al., 2018), publications applying his abundance-based multiple-site framework have been lagging. Moreover, estimation of local and species contributions to β-diversity is receiving increasing

scientific interest (e.g., Lopes et al., 2014; Tonkin et al., 2016; Heino & Grönroos, 2017; Vilmi et al., 2017). Nevertheless, to our knowledge, diatom studies on LCBD in terms of replacement and nestedness as well as the comparison of their incidence- and abundance-based results are absent.

The first aim of this study was to investigate the driving mechanisms of benthic diatom communities in small freshwater lakes of the Carpathian Basin: whether they are assembled merely due to the selection forces of the local environment or spatial variables are also important. Distances between our sampling sites can be considered as intermediate (2-400 km) and it covers regional scale instead of continental. Furthermore, environmental parameters vary reasonably across the sampled lakes (Table S1), however, none of them represents such extremely stressed environments as for instance, natural shallow saline lakes of the Carpathian Basin. Therefore, we assumed that both spatial distance between sites and local environmental characteristics should equally affect the development of diatom communities.

Our second goal was to estimate the regional β-diversity of diatom assemblages formed by metacommunity processes and to assess whether it is enhanced mainly by species turnover or nestedness related to the richness difference between sites. Similarly to the findings reported for most biota at low- or mid-latitude ecosystems (e.g., Tisseuil et al., 2012; Maloufi et al., 2016; Viana et al., 2016; Soininen et al., 2018; Szabó et al., 2018), we expected a high β-diversity of diatom communities due to the high degree of species turnover and a much smaller role of the nestedness component.

Furthermore, we intended to assess if sampled lakes contribute equally to β-diversity or some of them plays a particularly important role with its unique community composition and to determine which factors are responsible for the established patterns. Also, we wanted to examine this issue in terms of species turnover and nestedness, as well. We assumed that

112 sampling sites where one or more environmental parameters deviate considerably from the
113 average, thereby resulting in unique species combinations and/or low species richness
114 (Legendre, 2014), should have the largest contribution to β-diversity.

115 Finally, we wanted to quantify to what extent the individual species contribute to β-
116 diversity in the sampled region. We hypothesized that species that are characteristic of
117 restricted environmental conditions should affect overall β-diversity to the greatest extent.

118 Legendre (2014) suggested that the spatial distance among sampling sites should be
119 taken into account when choosing a dissimilarity index. Abundance-based calculations are
120 presumed to be appropriate at small spatial scales where species more likely differ in their
121 abundances rather than in their incidences. In contrast, incidence-based calculations are more
122 preferable within large spatial extents where sampling sites probably host different species.
123 Therefore, we aimed to test each of our hypotheses both with abundance- and incidence-based
124 analyses and to compare whether they provide distinct results.

125

126 **Material and Methods**

127 *Study sites, sampling and laboratory analyses*

128 In August 2010, a total of 38 freshwater lakes were sampled in the Carpathian Basin (Fig. 2,
129 Table S2). Each of them had a surface smaller than 3 km^2 and their altitude varied between 73
130 and 311 m (Table S2). Altitude of sampling sites (Table S1) were measured in Google Earth
131 Pro. The geographical distance between two sampling sites ranged from 2 to 400 km.

132 Phytobenthos samples were collected in the littoral region primarily from common
133 reed (*Phragmites australis* (Cav.) Trin. ex Steud.) or from other characteristic emergent
134 macrophytes, such as sedge (*Carex* sp.) or bulrush (*Typha* sp.) (CEN, 2003; King et al.,
135 2006). In each case, five macrophyte stems of the same species were chosen and starting ca. at
136 10 cm below the water surface, their 15-cm sections were cut. In some lakes, where

137 macrophyte vegetation was not characteristic or was absent, benthic diatoms were taken from
138 permanently-submerged natural stones, boughs or in case of their absence, from mud surface
139 with pipette. In each lake, only one type of substrates was sampled. Diatom valves were
140 cleaned by hot hydrogen-peroxide method and embedded in Zrax© resin (CEN, 2003).
141 Species were identified at 1000× magnification using Zeiss Axio Imager A1 with
142 Planapochromat DIC lense (Zeiss, 518N) according to the standard taxonomic guides (Bey &
143 Ector, 2010; Hofmann et al., 2011; Krammer, 2000, 2002, 2003; Krammer & Lange-Bertalot,
144 1991, 1999a, b, 2000; Lange-Bertalot, 2001; Lange-Bertalot et al., 2011; Levkov, 2009;
145 Levkov et al., 2013). In each sample, a minimum of 400 diatom valves was counted. All
146 diatom taxa (identified at species or genera level) were regarded as individual species and
147 were included in each subsequent statistical analysis.
148 Furthermore, water physical and chemical characteristics were determined for each
149 sampling site. Water temperature, oxygen saturation (DO%), conductivity, pH and turbidity
150 were measured *in situ* using an HQ40d Hach Lange multimeter. In laboratory, concentration
151 of HCO_3^-, Cl^-, SO_4^{2-} and COD were determined titrimetrically (APHA, 1998), whereas NO_2^-,
152 NO_3^-, NH_4^+, SRP, TP (APHA 1998) and SRSi (Wetzel & Likens, 2000)
153 spectrophotometrically.
154
155 *Statistical analyses*
156 Prior to the metacommunity-analyses, non-metric multidimensional scaling (NMDS) was
157 performed to visualize whether community composition of benthic diatoms was separated
158 according to the substrate types. NMDS was conducted based on the Hellinger-transformed
159 species abundance data applying Bray-Curtis distance. The NMDS projection displayed that
160 benthic diatom communities were not separated according to the substrate types and their

161 distribution was relatively homogeneous (Fig. S1). Therefore, all samples were included in
162 the subsequent statistical analyses.

163 The relative contribution of pure and shared effect of environmental heterogeneity and
164 spatial distance to variability of diatom communities was investigated with variation
165 partitioning method (Peres-Neto et al., 2006). In this analysis, two data matrices were used to
166 define the two explanatory variable groups. One of that was the group "environmental
167 heterogeneity", which consisted of the first two principal components' scores produced by a
168 principal component analysis (PCA) on a correlation matrix of standardized physical and
169 chemical parameters. In the group 'spatial distance', distance-based Moran's eigenvectors
170 (dbMEMs) were included as explanatory variables computed by principal coordinate analysis
171 (PCoA) of a truncated geographic distance matrix among sampling locations (Borcard &
172 Legendre, 2002; Borcard et al., 2004). Variation partitioning was performed both for
173 Hellinger transformed species abundance (Legendre & Gallagher, 2001; Borcard et al., 2011)
174 and species incidence data. ANOVA (permutations = 999) of RDA models were run to assess
175 the significance of adjusted R^2 values for testable fractions (pure environmental heterogeneity
176 and spatial distance).

177 To estimate overall β-diversity of diatom communities across the 38 sampling sites,
178 first we calculated abundance-based multiple-site Bray-Curtis dissimilarity (β_{BC}), which was
179 partitioned into its two components: abundance balanced variation ($\beta_{BC.BAL}$) and abundance
180 gradients ($\beta_{BC.GRA}$) (Baselga, 2017). Then, we transformed diatom abundance data into
181 presence-absence data and performed the same estimation using incidence-based multiple-site
182 Sørensen dissimilarity index (Baselga, 2010). Sørensen index (β_{SOR}) was also divided into its
183 components: turnover (β_{SIM}) and nestedness resultant (β_{NES}) component (Baselga et al., 2007;
184 Baselga, 2010).

Local contribution to β-diversity was calculated for each sampling site to quantify their ecological uniqueness. The computation was carried out both for abundance ($LCBD_{D\%diff}$) and presence-absence ($LCBD_{DS}$) data based on indices from the Baselga-family, Sørensen group. We used percentage different dissimilarity ($D\%_{diff}$) for quantitative (Baselga, 2013) and Sørensen dissimilarity (D_S) for binary data (Baselga, 2010). To stratify Euclidean property, we applied square-root transformation for dissimilarity matrices ($D\%_{diff}$ and $D_{S)}$ (Legendre & De Cáceres, 2013). To assess how unique each site is in terms of species replacement and nestedness, LCBD values were computed for replacement ($LCBD_{ReplB\%diff}$, $LCBD_{ReplBS}$) and nestedness ($LCBD_{NesB\%diff}$, $LCBD_{NesBS}$) decomposing $LCBD_{D\%diff}$ and $LCBD_{DS}$ (Legendre & De Cáceres, 2013).

To describe the relative importance of individual species in affecting overall β-diversity, we calculated species contribution to β-diversity for Hellinger-transformed abundance data ($SCBD_{ab}$) and for species incidence data ($SCBD_{pa}$) (Legendre & De Cáceres, 2013).

Since LCBD and SCBD indices (response variables) exhibit relative contribution data taking values between 0 and 1, generalized additive models (GAMs) using beta regression family with logit link function (Wood et al., 2016) were applied to investigate the relationship of $LCBD_{D\%diff}$, and $LCBD_{DS}$ with the local species richness as well as the relationship of $SCBD_{ab}$ and $SCBD_{pa}$ with the number of sites occupied by a given species and with the total abundance of the species. We run regression tree model analyses (Breiman et al., 1984) to find the most important environmental factors determining the variation in LCBD indices ($LCBD_{D\%diff}$, $LCBD_{DS}$, $LCBD_{ReplB\%diff}$, $LCBD_{ReplBS}$, $LCBD_{ReplB\%diff}$, $LCBD_{ReplBS}$). Furthermore, Pearson correlation coefficient was computed for each pair of LCBD indices to estimate the correlation between them.

R statistical software (R. 3.4.1; R Development Core Team, 2017) was used to conduct statistical analyses. We applied codep (Guenard et al., 2017) and ape (Paradis et al., 2004) R packages for dbMEM analysis and PCoA, and vegan (Oksanen et al., 2017) for variation partitioning. Multiple-site β-diversity indices were calculated in betapart (Baselga et al., 2017), LCBD and SCBD indices in adespatial (Dray et al., 2017), ade4 (Dray & Dufour, 2007) R packages and with beta.div function (Legendre & De Cáceres, 2013). Regression tree model analyses and GAMs were conducted and illustrated using rpart (Therneau et al., 2017), rpart.plot (Milborrow, 2017), mgcv (Wood, 2011) and ggplot2 (Wickham, 2009) R packages.

Results

Physical and chemical parameters varied considerably among the 38 lakes, many of them had a higher standard deviation than the mean (Table S1). According to the PCA results (Fig. 3), 33.7% of the variance in environmental factors is explained by PC1 axis and 17.3% by PC2 axis. In descending order, HCO_3^-, conductivity, TP and SRP showed the highest correlation with PC1 axis (absolute values of Pearson correlation coefficients were above 0.8) and had the highest PC1 loading. Variables correlated most with PC2 axis (absolute values of Pearson correlation coefficients were above 0.6) and possessing the highest PC2 loading were O_2 saturation and pH.

Based on the results of variation partitioning (Fig. 4), the establishment of diatom community composition was affected significantly by environmental heterogeneity and spatial distance among the sampling sites as well. However, either in case of species abundance or incidence data, the pure spatial distance explained a slightly higher proportion (7.3% and 3.4%) of community variation than environmental heterogeneity alone (2.8% and 2.1%).

In the 38 phytobenthos samples, 273 diatom taxa were found, of which 269 were identified at species level and four at genus level. The number of species showed high

234 variability: its lowest value was 20 and the highest was 66 (average and standard deviation: 42
235 ± 12). We found high overall β-diversity of diatom communities according to the abundance-
236 based (β_{BC}=0.956) as well as the incidence-based (β_{SOR}=0.934) multiple-site framework. In
237 both cases, β-diversity patterns were enhanced mainly by the component accounting for
238 species substitution (abundance balanced variation: $\beta_{BC.BAL}$=0.953 and turnover: β_{SIM}=0.914)
239 whereas the component accounting for subsets (abundance gradients: $\beta_{BC.GRA}$=0.003 and
240 nestedness β_{NES}=0.020) was very low.

241 We found strong positive correlation between $LCBD_{D\%diff}$ and $LCBD_{ReplB\%diff}$ as well
242 as between $LCBD_{DS}$ and $LCBD_{ReplBS}$ (Pearson correlation coefficients were 0.98 and 0.94,
243 respectively) furthermore, $LCBD_{NesB\%diff}$ correlated negatively with $LCBD_{D\%diff}$ and
244 $LCBD_{ReplB\%diff}$ (Pearson correlation coefficients were -0.51 and -0.47, respectively). For any
245 other pairs of indices, no significant correlation was displayed (Table S3). GAMs and
246 regression tree model analyses revealed that distinct factors affect the LCBD indices using
247 abundance and incidence data. There was no significant relationship between $LCBD_{D\%diff}$ and
248 local species richness, but $LCBD_{DS}$ showed a significant decrease with the increase of species
249 richness (Table 1, Fig. 5a-b).

250 Sites with the highest local contribution to β-diversity were different when conducting
251 computations on species abundance and presence-absence matrix. These two types of data
252 revealed different results also during the investigation of sampling sites' uniqueness in terms
253 of species replacement and nestedness. Sites possessing the highest $LCBD_{D\%diff}$ index
254 (>0.030) were CSA, CSI, HAR, ÖRE and SZT (Fig. S2a), and according to the regression tree
255 model analyses, environmental variables driving $LCBD_{D\%diff}$ were TP and NO_3^- (Fig. 6a).
256 Similarly, sampling sites with the highest $LCBD_{ReplB\%diff}$ value (>0.034) were CSA, CSI,
257 HAR, ÖRE and SZT (Fig. S2b) where SRP and COD were the most decisive (Fig. 6b). In
258 turn, sites represented by the highest $LCBD_{NesB\%diff}$ (>0.115) were KHT, TDO, VDK and

MÁM (Fig. S2c) determined primarily by TP, COD and pH (Fig. 6c). In case of the incidence-based data, the highest $LCBD_{DS}$ indices (>0.031) were found at sites HÁM, KEN, KFT, SÁR, VAD and PIR (Fig. S3a), where SRP, COD and SRSi had the most important effect (Fig. 7a). HÁM, KEN, KFT, SÁR and VAD (Fig. S3b) achieved the highest $LCBD_{ReplBS}$ value (>0.036) affected mainly by SRP and NH_4^+ (Fig. 7b). Sites with outstanding $LCBD_{NesBS}$ index (>0.095) were PIR, TÚR and TOL (Fig. S3c) driven by SRP and TP concentration (Fig. 7c).

Contribution of the individual species to β-diversity depended on the type of the applied data matrix (abundance- or incidence-based). According to the GAMs' results, SCBD using abundance data ($SCBD_{ab}$) depended both on the number of sites occupied by the given species and on the total abundance of the species (Table 1, Fig. 5c, e): it showed an increasing trend with the increase of both explanatory variables. In turn, SCBD based on incidence data ($SCBD_{pa}$) was significantly related only to the number of occupied sites and a unimodal (hump-shaped) relationship was revealed between them (Table 1, Fig. 5d, f): $SCBD_{pa}$ increased up to 20 occupied sites and then, it started to decrease. Species with the highest $SCBD_{ab}$ value (>0.05) were *Achnanthidium minutissimum* (Kützing) Czarnecki, *Amphora pediculus* (Kützing) Grunow and *Cocconeis placentula* Ehrenberg, all of which occupied high number of samples (≥28) and were present with high total abundance (≥1060 individuals counted during the study). In contrast, *Eolimna minima* (Grunow) Lange-Bertalot, *Halamphora veneta* (Kützing) Levkov, *Nitzschia palea* var. *tenuirostris* Grunow, *Nitzschia palea* var. *debilis* (Kützing) Grunow and *Nitzschia supralitorea* Lange-Bertalot had the highest $SCBD_{pa}$ (>0.01). These species occurred at intermediate proportion of sites (at 18-20 sites) and with moderate total abundance (110-354 individuals).

Discussion

Structuring drivers and β-diversity of diatom communities

In accordance with our first hypothesis, the composition of benthic diatom communities in the studied small, freshwater lakes of the Carpathian Basin depended significantly on the spatial variables, however, the filtering effect of the lakes' local environmental characteristics played also a significant role. Studies using variation partitioning to unravel metacommunity mechanisms assume, in general, i) species-sorting if solely the "environmental variables" fraction explains significantly the community structures; ii) neutral theory or patch dynamics if only the "spatial variables" fraction is significant and iii) mass-effect concept or the combination of species-sorting and mass-effect if both fractions have significant explanatory power (Cottenie, 2005; Soininen, 2014). However, instead of regarding metacommunity concepts as distinct alternatives, considering them as continuum is suggested (Alonso et al., 2006; Gravel et al., 2006; Leibold & McPeek, 2006; Adler et al., 2007; Chase, 2007). It is impossible to firmly determine the boundaries between the types of metacommunities due to several interfering factors (Leibold & Chase, 2017). The degree of environmental heterogeneity within the studied area and the traits of species, such as size and dispersal rate, greatly influence the response of species to habitat heterogeneity. Relatively large species with low dispersal rates are assumed to be structured according to spatial variables due to their limited dispersion complying with the neutral theory and patch dynamics rather than by environmental characteristics. In contrast, smaller species with better dispersion abilities are likely driven by habitat heterogeneity because they might be able to respond more sensitively even to the minor environmental differences (Hájek et al., 2011; De Bie et al., 2012; Heino, 2013). In case of intermediate dispersal rates, dispersion limitation is not probable and environmental heterogeneity inherent to species-sorting mechanisms is the most decisive, whereas structure of the best dispersing species is slightly better explained by the spatial variables and habitat heterogeneity is less important that is, mass-effect will become prevalent

309 (Leibold & Chase, 2017). Our variation partitioning results, and taking into account the small
310 size and the effective passive dispersion (Kristiansen, 1996; Finlay, 2002) of diatoms, point to
311 the fact that at intermediate spatial scale in the Carpathian Basin lake benthic diatoms were
312 assembled in conformity with the mass-effect theory. However, despite that diatoms are
313 regarded as relatively well dispersing organisms within large areas (e.g., at continental or
314 global scale), geographic separation tend to limit their ubiquitous dispersal thus showing pure
315 spatial patterns, which can be explained by the neutral theory (Heino et al., 2010).
316 Nevertheless, it would be difficult to decide exclusively for one metacommunity concept
317 without quantifying accurately the species' dispersal rate and the strength of environmental
318 gradients within the studied region (Logue et al., 2011; Lindström & Langenheder, 2012;
319 Maloufi et al., 2016). In addition, the observed high proportion of unexplained variation
320 (residuals) probably deriving from unmeasured environmental parameters, undersampling of
321 rare species and stochastic processes should not be ignored during the interpretation of the
322 observed patterns. For instance, if an originally unmeasured variable were spatially structured,
323 the importance of the "spatial variables" fraction would increase, whereas if it were not
324 spatially structured, residuals would be higher, leading to distinct conclusions regarding
325 metacommunity theories (Leibold & Chase, 2017). Moreover, unregulated ecological drifts
326 and colonization-extinction stochasticity (predicted by the neutral theory and patch dynamics,
327 respectively) might also increase residual variation (Hubbell, 2001; Vellend, 2010, 2016;
328 Leibold & Chase, 2017).

329 We experienced a high average of local diatom species richness, which confirms
330 Mouquet & Loreau's (2003) theory that consequent on mass-effect, α-diversity should
331 increase if dispersal rate slightly increases. However, this process should result in a decreased
332 β-diversity among sites. Contrary to this, but in agreement with our expectations and previous
333 findings (Tisseuil et al., 2012; Maloufi et al., 2016; Viana et al., 2016; Soininen et al., 2018;

Szabó et al., 2018) that at mid-latitudes (like the Carpathian Basin) driving mechanisms expounded above, resulted in very high β-diversity primarily due to the high degree of species turnover among the sampling sites. In turn, nestedness resulted from richness differences was inconsiderable based on our analyses. In the meta-analysis by Soininen et al. (2018) species turnover and total β-diversity showed strong correlation as both quantify the compositional dissimilarities between samples, whereas nestedness is represented with several times smaller proportion (even close to zero) than turnover and it may only measure the bias caused by richness differences. They also described that β-diversity and its turnover component are slightly smaller near the poles, which could be explained by the more homogeneous environment, less limited species dispersion (Mouquet & Loreau, 2003; Leibold et al., 2004) and less pronounced biotic interactions (Willig et al., 2003; Schemske et al., 2009). Towards higher latitudes, where glaciation might have played an important role in the local and regional extinction and recolonization processes, the increase of nestedness was found (Soininen et al., 2018). Either species abundance or presence-absence data were applied during the analyses, we were able to draw the same conclusion that both the local environment and the spatial distance influenced the benthic diatom assemblages and high β-diversity was enhanced by species turnover. However, similarly to previous studies (Heino et al., 2010; Vilmi et al., 2016; Szabó et al., 2018), the unexplained variation in community structure was higher when only the incidence of diatom species was considered.

Local contribution of sampling sites to β-diversity

Calculation of LCBD is suitable for quantifying which sites contribute more (or less) to β-diversity than the mean and thereby for evaluating the ecological uniqueness of communities at each sites (Legendre & De Cáceres, 2013). Local contribution to β-diversity and local contribution in terms of species replacement showed a strong positive relationship applying

359 either abundance- or incidence-based data. However, in case of using abundance data, LCBD
360 for nestedness decreased significantly with increasing LCBD and LCBD for replacement.
361 Accordingly, sites with highest uniqueness in terms of replacement contributed to the greatest
362 extent to total β-diversity of diatom communities, as well. This may be related to the fact that
363 in general, total β-diversity also correlates positively with its turnover component and
364 negatively with its nestedness component (Soininen et al., 2018). It is supposed that species-
365 rich sites exhibit low LCBD due to the greater chance of sharing species with other
366 communities (Maloufi et al., 2016). Nevertheless, our assumption that sites with low diatom
367 species richness have greater contribution to the regional β-diversity than sites with higher
368 richness, was only partly confirmed by the results. The declining trend in LCBD with
369 increasing local richness was observed both for abundance and presence-absence data, but the
370 relationship was significant only for species incidences. A part of former studies confirms,
371 whereas some of them contradicts our findings depending on the organisms and the habitat
372 type targeted. Applying abundance data for stream (Vilmi et al., 2017) and pond (Teittinen et
373 al., 2017) diatom communities, negative correlation between LCBD and species richness was
374 reported, however, this relationship was not evident for lake benthic diatoms (Vilmi et al.,
375 2017). In case of dung beetles (Da Silva & Hernández, 2014) and stream insect assemblages
376 (Heino & Grönroos, 2017), LCBD decreased significantly with increasing local species
377 richness if calculations were conducted on presence-absence data, which is in line with our
378 findings. Consequently, we concluded that sites sustaining less diverse communities have
379 greater ecological uniqueness, however, this coherence varies among different groups of
380 organisms and ecosystems, furthermore also depends largely on the data type applied.
381 Our results revealed that local environmental variables affected sampling sites'
382 contribution to β-diversity, including its extension to replacement and nestedness, as well.
383 Although sites with highest LCBD indices were different based on abundance- and incidence-

based community data, we did not find explicit contrast in their main driving variables. Most decisive factors were phosphorus forms for each LCBD index, which corroborates our hypothesis, since these parameters displayed relatively high variance among the sites. Additionally, nitrogen forms, pH, COD and SRSi were also crucial in evolving sites' ecological uniqueness for diatom communities. These findings are not surprising, since nutrient supply plays a key role in establishment of autotrophic algal assemblages and trophic status is also related, for instance, to oxygen conditions and pH. Thereby, it affects indirectly the physiological processes of aquatic organisms (Soininen, 2007). The above chemical parameters have already been emphasized as master variables for freshwater lake diatom communities in several previous studies (e.g., Hall & Smol, 1992; King et al., 2000; Lim et al., 2001; Soininen, 2007). In addition, pH was found as one of the most influential variables for subarctic ponds' contribution to β-diversity of diatom communities (Teittinen et al., 2017). In turn, some publications targeting β-diversity assessments reported that LCBD was not well determined by local environmental characteristics, for instance, in case of stream insects (Heino & Grönroos, 2017) and invertebrates (Tonkin et al., 2016).

Species contribution to β-diversity

With respect to species contribution to β-diversity, results published for different biota and ecosystems are congruent, however, abundance- and incidence-based calculations displayed fundamentally distinct patterns similarly to our findings. Gaston et al. (2006) emphasized the tight link between abundance, its spatial variation and the number of occupied sites by a given species, which may be related to our observations that diatom species occupying a high number of lakes and represented by high abundance contributed the most to overall β-diversity. That is, contrary to our hypothesis, common diatom species such as *Achnanthidium minutissimum*, *Amphora pediculus* and *Cocconeis placentula* with extensive ecological

amplitude (Hofmann et al., 2011) and variable abundance at different sites exerted the greatest impact on β-diversity. This pattern prevailed only in case of abundance-based SCBD similarly to observations by Heino & Grönroos (2017) for stream insects and by Vilmi et al. (2017) for stream and lake diatom communities. Our incidence-based calculations revealed that species with intermediate occupancy had the largest contribution to β-diversity, which was also observed by Heino & Grönroos (2017). This may be due to the fact that occupancy of these species can vary largely across the sites (Gaston et al., 2006). Species with the highest incidence-based SCBD were *Eolimna minima*, *Halamphora veneta*, *Nitzschia palea* var. *tenuirostris*, *N. palea* var. *debilis* and *N. supralitorea*, which are relatively common and possess intermediate-sized niches (Hofmann et al., 2011). Also, their total abundance was intermediate in our data set but in this case, the relationship between SCBD and species' abundance was statistically not significant. However, it is important to note that both dependent (SCBD index) and explanatory variables (occupancy and species' total abundance) of the models are not independent mathematically, since each of them is conducted from the same raw community data (even abundance or presence-absence), which might have affected the strong relationship between them (Legendre & De Cáceres, 2013; Heino & Grönroos, 2017).

Conclusions

At intermediate spatial scale (2-400 km) of a mid-latitude region, where physical and chemical parameters across small freshwater lakes are relatively, but not extremely heterogeneous, benthic diatoms were assembled conforming most to the mass-effect metacommunity concept. However, because patterns are largely dependent on several factors (such as scale of heterogeneity, environmental variables considered during the study, dispersal rates, size of species pool and stochastic processes), conclusions should be drawn with

caution. The high α-diversity (average of local species richness) found in the region, is in line with the mass-effect paradigm, which is, in turn, inconsistent with the high β-diversity enhanced mainly by species turnover. Freshwater lakes in the Carpathian Basin with the highest contribution to overall β-diversity (and with the highest ecological uniqueness in terms of turnover, too) hosted a lower number of diatom species than the average, however, biodiversity conservation, in general, focuses on preserving species-rich sites. Furthermore, β-diversity was related mainly to the regionally common species that have medium-sized or broad niches, instead of the ecologically restricted ones. Therefore, we advocate the previous suggestions made by Heino & Grönroos (2017) and Vilmi et al. (2017) that if a study aims comprehensive conservation planning, a simultaneous application of LCBD and SCBD indices combining with the focus on species-rich ecosystems and rare species would be sufficient. Moreover, although abundance-based and incidence-based analyses led us to the same conclusions regarding metacommunity concept and sites' ecological uniqueness, they displayed different patterns of SCBD. Consequently, for assessing species' ecological uniqueness during an extensive research of metacommunities, we recommend conducting the analyses both on species abundance and binary data, especially in case of conservation objectives.

Acknowledgements

We thank the colleagues of the Department of Limnology for their contribution in field sampling and in laboratory work. The study was supported by the Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015 and the National Research Development and Innovation Office (NKFIH K120595). Beáta Szabó was supported by the ÚNKP-17-3-IV-PE-5 New National Excellence Program of the Ministry of Human Capacities.

Conflict of Interest

The authors declare that they have no conflict of interest.

462 **References**

463 Adler, P. B., J. HilleRislambers & J. M. Levine, 2007. A niche for neutrality. Ecology Letters
464 10: 95–104.

465 Alonso, D., R. S. Etienne & A. J. McKane, 2006. The merits of neutral theory. Trends in
466 Ecology and Evolution 21: 451–457.

467 APHA (American Public Health Association), 1998. Standard methods for the examination of
468 water and wastewater, 20th edition. United Book Press, Maltimore (MD).

469 Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity.
470 Global Ecology and Biogeography 19: 134–143.

471 Baselga, A., 2013. Separating the two components of abundance-based dissimilarity: Balanced
472 changes in abundance vs. abundance gradients. Methods in Ecology and Evolution 4:
473 552–557.

474 Baselga, A., 2017. Partitioning abundance-based multiple-site dissimilarity into components:
475 balanced variation in abundance and abundance gradients. Methods in Ecology and
476 Evolution 8: 799–808.

477 Baselga, A., A. Jiménez-Valverde & G. Niccolini, 2007. A multiple-site similarity measure
478 independent of richness. Biology Letters 3: 642–645.

479 Baselga, A., D. Orme, S. Villeger, J. De Bortoli & F. Leprieur, 2017. betapart: Partitioning
480 Beta Diversity into Turnover and Nestedness Components. R package version 1.4-1.
481 https://CRAN.R-project.org/package=betapart

482 Bey, M-Y. & L. Ector, 2010. Atlas des diatomées des cours d'eau de la région Rhône-Alpes,
483 Tome 1-6. Direction régionale de l'Environnement, de l'Aménagement et du Logement
484 Rhône-Alpes, Lyon.

485 Borcard, D. & P. Legendre, 2002. All-scale spatial analysis of ecological data by means of
486 principal coordinates of neighbour matrices. Ecological Modelling 153: 51–68.

487 Borcard, D., F. Gillet & P. Legendre, 2011. Numerical Ecology with R. Springer, New York.
488 Borcard, D., P. Legendre, C. Avois-Jacquet & H. Tuomisto, 2004. Dissecting the spatial
489 structure of ecological data at multiple scales. Ecology 85: 1826–1832.
490 Breiman, L., J. H. Friedman, R. A. Olshen & C. J. Stone, 1984. Classification and Regression
491 Trees. Wadsworth International Group, Belmont.
492 CEN (Comité Européen de Normalisation), 2003. Water Quality Guidance Standard for the
493 Routine Sampling and Pretreatment of Benthic Diatoms from Rivers. EN 13946:2003,
494 Geneva.
495 Chase, J. M., 2007. Drought mediates the importance of stochastic community assembly.
496 Proceedings of the National Academy of Sciences 104: 17430–17434.
497 Conradi, T., V. M. Temperton & J. Kollmann, 2017. Beta diversity of plant species in human-
498 transformed landscapes: Control of community assembly by regional productivity and
499 historical connectivity. Perspectives in Plant Ecology, Evolution and Systematics 24: 1–
500 10.
501 Cottenie, K., 2005. Integrating environmental and spatial processes in ecological community
502 dynamics. Ecology Letters 8: 1175–1182.
503 Da Silva, P. G. & M. I. M. Hernández, 2014. Local and regional effects on community
504 structure of dung beetles in a mainland-island scenario. PLoS One 9: e111883.
505 De Bie, T., L. De Meester, L. Brendonck, K. Martens, B. Goddeeris, D. Ercken, H. Hampel, L.
506 Denys, L. Vanhecke, K. Van der Gucht, J. Van Wichelen, W. Vyverman & S. A. J.
507 Declerck, 2012. Body size and dispersal mode as key traits determining metacommunity
508 structure of aquatic organisms. Ecology Letters 15: 740–747.
509 Dong, X., B. Li, F. He, Y. Gu, M. Sun, H. Zhang, L. Tan, W. Xiao, S. Liu & Q. Cai, 2016.
510 Flow directionality, mountain barriers and functional traits determine diatom
511 metacommunity structuring of high mountain streams. Scientific Reports 6: 24711.

512 Dray, S. & A. B. Dufour, 2007. The ade4 package: implementing the duality diagram for
513 ecologists. Journal of Statistical Software. 22: 1–20.
514 Dray, S., G. Blanchet, D. Borcard, G. Guenard, T. Jombart, G. Larocque, P. Legendre, N.
515 Madi & H. H. Wagner, 2017. adespatial: Multivariate Multiscale Spatial Analysis. R
516 package version 0.0-8. https://CRAN.R-project.org/package=adespatial
517 Finlay, B. J., 2002. Global dispersal of free-living microbial eukaryote species. Science 296:
518 1061–1063.
519 Gaston, K. J., P. A. V. Borges, F. He & C. Gaspar, 2006. Abundance, spatial variance and
520 occupancy: Arthropod species distribution in the Azores. Journal of Animal Ecology 75:
521 646–656.
522 Göthe, E., D. G. Angeler, S. Gottschalk, S. Löfgren & L. Sandin, 2013. The influence of
523 environmental, biotic and spatial factors on diatom metacommunity structure in Swedish
524 headwater streams. PLoS One 8: e72237.
525 Gravel, D., C. D. Canham, M. Beaudet & C. Messier, 2006. Reconciling niche and neutrality:
526 The continuum hypothesis. Ecology Letters 9: 399–409.
527 Guenard, G., P. Legendre & with contributions from B. Pages, 2017. codep: Multiscale
528 Codependence Analysis. R package version 0.6-5. https://CRAN.R-
529 project.org/package=codep
530 Hájekm, M., J. Roleček, K. Cottenie, K. Kintrová, M. Horsák, A. Poulíčková, P. Hájková, M.
531 Fránková & D. Dítě, 2011. Environmental and spatial controls of biotic assemblages in a
532 discrete semi-terrestrial habitat: Comparison of organisms with different dispersal
533 abilities sampled in the same plots. Journal of Biogeography 38: 1683–1693.
534 Hall, R. I. & J. P. Smol, 1992. A weighted—averaging regression and calibration model for
535 inferring total phosphorus concentration from diatoms in British Columbia (Canada)
536 lakes. Freshwater Biology 27: 417–434.

Heino, J., 2013. Environmental heterogeneity, dispersal mode, and co-occurrence in stream macroinvertebrates. Ecology and Evolution 3: 344–355.

Heino, J., L. M. Bini, S. M. Karjalainen, H. Mykrä, J. Soininen, L. C. G. Vieira & J. A. F. Diniz-Filho, 2010. Geographical patterns of micro-organismal community structure: Are diatoms ubiquitously distributed across boreal streams? Oikos 119: 129–137.

Heino, J. & M. Grönroos, 2017. Exploring species and site contributions to beta diversity in stream insect assemblages. Oecologia 183: 151–160.

Hofmann, G., M. Werum & H. Lange-Bertalot, 2011. Diatomeen im Süßwasser- Benthos von Mitteleuropa. Koeltz Scientific Books, Königstein.

Hubbell, S. P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton, New Jersey.

King, L., P. Barker & R. I. Jones, 2000. Epilithic algal communities and their relationship to environmental variables in lakes of the English Lake District. Freshwater Biology 45: 425–442.

King, L., G. Clarke, H. Bennion, M. Kelly & M. Yallop, 2006. Recommendations for sampling littoral diatoms in lakes for ecological status assessments. Journal of Applied Phycology 18: 15–25.

Krammer, K., 2000. Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats (Vol. 1. The Genus *Pinnularia*). ARG Gantner Verlag KG, Ruggel.

Krammer, K., 2002. Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats (Vol. 3. *Cymbella*). ARG Gantner Verlag KG, Ruggel.

Krammer, K., 2003. Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats (Vol. 4. *Cymbopleura*, *Delicata*, *Navicymbula*, *Gomphocymbellopsis*, *Afrocymbella*). ARG Gantner Verlag KG, Ruggel.

561 Krammer, K. & H. Lange-Bertalot, 1991. Bacillariophyceae 4. Teil: Achnanthaceae. Kritische
562 Ergänzungen zu *Navicula* (Lineolatae) und *Gomphonema*. In Ettl, H., J. Gerloff, H.
563 Heynig & D. Mollenhauer D (eds), Süsswasserflora von Mitteleuropa, Band 2/4.
564 Spektrum Akademischer Verlag, Heidelberg.
565 Krammer, K. & H. Lange-Bertalot, 1999a. Bacillariophyceae 1. Teil: Naviculaceae. In Ettl,
566 H., J. Gerloff, H. Heynig & D. Mollenhauer D (eds), Süsswasserflora von Mitteleuropa,
567 Band 2/1. Spektrum Akademischer Verlag, Heidelberg.
568 Krammer, K. & H. Lange-Bertalot, 1999b. Bacillariophyceae 2. Teil: Bacillariaceae,
569 Epithemiaceae, Surirellaceae. In Ettl, H., J. Gerloff, H. Heynig & D. Mollenhauer D
570 (eds), Süsswasserflora von Mitteleuropa, Band 2/2. Spektrum Akademischer Verlag,
571 Heidelberg.
572 Krammer, K. & H. Lange-Bertalot, 2000. Bacillariophyceae 3. Teil: Centrales, Fragilariaceae,
573 Eunotiaceae. In Ettl, H., J. Gerloff, H. Heynig & D. Mollenhauer D (eds),
574 Süsswasserflora von Mitteleuropa, Band 2/3, Spektrum Akademischer Verlag,
575 Heidelberg.
576 Kristiansen, J., 1996. Dispersal of freshwater algae — a review. Hydrobiologia 336: 151–157.
577 Lange-Bertalot, H., 2001. Diatoms of Europe Diatoms of the European Inland Waters and
578 Comparable Habitats (Vol. 2. *Navicula* sensu stricto. 10 genera separated from *Navicula*
579 sensu lato. *Frustulia*). ARG Gantner Verlag KG, Ruggel.
580 Lange-Bertalot, H., M. Malgorzata & A. Witkowski, 2011. Diatoms of Europe: Diatoms of the
581 European Inland Waters and Comparable Habitats (Vol. 6. *Eunotia* and some related
582 genera). ARG Gantner Verlag KG, Ruggel.
583 Langenheder, S. & H. Ragnarsson, 2007. The role of environmental and spatial factors for the
584 composition of aquatic bacterial communities. Ecology 88: 2154–2161.

585 Legendre, P., 2014. Interpreting the replacement and richness difference components of beta
586 diversity. Global Ecology and Biogeography 23: 1324–1334.
587 Legendre, P. & M. De Cáceres, 2013. Beta diversity as the variance of community data:
588 Dissimilarity coefficients and partitioning. Ecology Letters 16: 951–963.
589 Legendre, P. & E. D. Gallagher, 2001. Ecologically meaningful transformations for ordination
590 of species data. Oecologia 129: 271–280.
591 Leibold, M. A. & J. M. Chase, 2017 Metacommunity Ecology, Princeton University Press,
592 Princeton, New Jersey.
593 Leibold, M. A. & M. A. McPeek, 2006. Coexistence of the niche and neutral perspectives in
594 community ecology. Ecology 87: 1399–1410.
595 Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D.
596 Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau & A. Gonzales, 2004. The
597 metacommunity concept: A framework for multi-scale community ecology. Ecology
598 Letters 7: 601–613.
599 Levkov, Z., 2009. Diatoms of Europe: Diatoms of the European Inland Waters and
600 Comparable Habitats (Vol. 5. *Amphora* sensu lato). ARG Gantner Verlag KG, Ruggel.
601 Levkov, Z., D. Metzeltin & A. Pavlov, 2013. Diatoms of Europe: Diatoms of the European
602 Inland Waters and Comparable Habitats (Vol. 7. *Luticola* and *Luticolopsis*). Koeltz
603 Scientific Books, Königstein.
604 Lim. D. S. S., C. Kwan, M. S. V. Douglas, 2001. Periphytic diatom assemblages from Bathurst
605 Island, Nunavut, Canadian High Arctic: An examination of community relationships and
606 habitat preferences. Journal of Phycology 37: 379–392.
607 Lindström, E. S. & S. Langenheder, 2012. Local and regional factors influencing bacterial
608 community assembly. Environmental Microbiology Reports 4: 1–9.

609 Logue, J. B., N. Mouquet, H. Peter & H. Hillebrand, 2011. Empirical approaches to
610 metacommunities: a review and comparison with theory. Trends in Ecology and
611 Evolution 26: 482–491.
612 Lopes, P. M., L. M. Bini, S. A. J. Declerck, V. F. Farjalla, L. C. G. Vieira, C. C. Bonecker, F.
613 A. Lansac-Toha, F. A. Esteves & R. L. Bozelli, 2014. Correlates of zooplankton beta
614 diversity in tropical lake systems. PLoS One 9: e109581.
615 Maloufi, S., A. Catherine, D. Mouillot, C. Louvard, A. Couté, C. Bernard & M. Troussellier,
616 2016. Environmental heterogeneity among lakes promotes hyper β-diversity across
617 phytoplankton communities. Freshwater Biology 61: 633–645.
618 Milborrow, S., 2017. rpart.plot: Plot 'rpart' Models: An Enhanced Version of 'plot.rpart'. R
619 package version 2.1.2. https://CRAN.R-project.org/package=rpart.plot
620 Mouquet, N. & M. Loreau, 2003. Community Patterns in Source-Sink Metacommunities. The
621 American Naturalist 162: 544–557.
622 Mykrä, H., J. Heino & T. Muotka, 2007. Scale-related patterns in the spatial and
623 environmental components of stream macroinvertebrate assemblage variation. Global
624 Ecology and Biogeography 16: 149–159.
625 Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin,
626 R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs & H. Wagner,
627 2017. vegan: Community Ecology Package. R package version 2.4-3. https://CRAN.R-
628 project.org/package=vegan
629 Paradis, E., J. Claude & K. Strimmer, 2004. APE: analyses of phylogenetics and evolution in
630 R language. Bioinformatics 20: 289–290.
631 Peres-Neto, P. R., P. Legendre, S. Dray & D. Borcard, 2006. Variation partitioning of species
632 data matrices: Estimation and comparison of fractions. Ecology 87: 2614–2625.

633 R Development Core Team, 2017. R: A Language and Environment for Statistical Computing.
634 R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0. http://www.R-
635 project.org
636 Schemske, D.W., G. G. Mittelbach, H. V. Cornell, J. M. Sobel & K. Roy, 2009. Is there a
637 latitudinal gradient in the importance of biotic interactions? Annual Review of Ecology,
638 Evolution, and Systematics 40: 245–269.
639 Soininen, J., 2007. Environmental and spatial control of freshwater diatoms - A review.
640 Diatom Research 22: 473–490.
641 Soininen, J., 2014. A quantitative analysis of species sorting across organisms and ecosystems.
642 Ecology 95: 3284–3292.
643 Soininen, J., J. Heino & J. Wang, 2018. A meta-analysis of nestedness and turnover
644 components of beta diversity across organisms and ecosystems. Global Ecology and
645 Biogeography 27: 96–109.
646 Szabó, B., E. Lengyel, J. Padisák, M. Vass, C. Stenger-Kovács, 2018. Structuring forces and
647 β-diversity of benthic diatom metacommunities in soda pans of the Carpathian Basin.
648 European Journal Phycology 53: 219-229.
649 Teittinen, A., J. Wang, S. Strömgård & J. Soininen, 2017. Local and geographical factors
650 jointly drive elevational patterns in three microbial groups across subarctic ponds. Global
651 Ecology and Biogeography 26: 973–982.
652 Therneau, T., B. Atkinson & B. Ripley, 2017. rpart: Recursive Partitioning and Regression
653 Trees. R package version 4.1-11. https://CRAN.R-project.org/package=rpart
654 Tisseuil, C., F. Leprieur, G. Grenouillet, M. Vrac & S. Lek, 2012. Projected impacts of climate
655 change on spatio-temporal patterns of freshwater fish beta diversity: A deconstructing
656 approach. Global Ecology and Biogeography 21: 1213–1222.

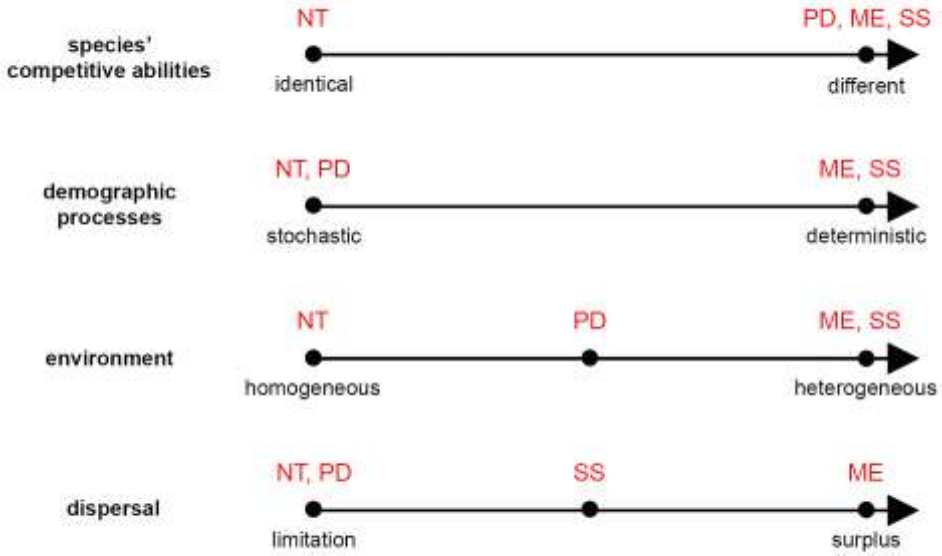
657 Tonkin, J. D., J. Heino, A. Sundermann, P. Haase & S. C. Jähnig, 2016. Context dependency
658 in biodiversity patterns of central German stream metacommunities. Freshwater Biology
659 61: 607–620.
660 Vellend, M., 2010. Conceptual Synthesis in Community Ecology. The Quarterly Review of
661 Biology 85: 183–206.
662 Vellend, M., 2016. The Theory of Ecological Communities (MPB-57). Princeton University
663 Press, New Jersey.
664 Viana, D. S., J. Figuerola, K. Schwenk, M. Manca, A. Hobæk, M. Mjelde, C. D. Preston, R. J.
665 Gornall, J. M. Croft, R. A. King, A. J. Green & L. Santamaría, 2016. Assembly
666 mechanisms determining high species turnover in aquatic communities over regional and
667 continental scales. Ecography (Cop) 39: 281–288.
668 Vilmi, A., S. M. Karjalainen & J. Heino, 2017. Ecological uniqueness of stream and lake
669 diatom communities shows different macroecological patterns. Diversity and
670 Distributions 23: 1042–1053.
671 Vilmi, A., S. M. Karjalainen, S. Hellsten & J. Heino, 2016. Bioassessment in a
672 metacommunity context: Are diatom communities structured solely by species sorting?
673 Ecological Indicators 62: 86–94.
674 Wetzel, R. G. & G. E. Likens, 2000 Limnological Analyses. Springer-Verlag, New York.
675 Wickham, H., 2009 ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
676 Willig, M. R., D. M. Kaufman, R. D. Stevens, 2003. Latitudinal gradients of biodiversity:
677 patterns, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics 34:
678 273–309.
679 Wood, S. N., 2011. Fast stable restricted maximum likelihood and marginal likelihood
680 estimation of semiparametric generalized linear models. Journal of the Royal Statistical
681 Society. Series B: Statistical Methodology 73: 3–36.

682 Wood, S. N., N. Pya & B. Säfken, 2016. Smoothing Parameter and Model Selection for
683 General Smooth Models. Journal of the American Statistical Association 111: 1548–
684 1563.
685

686 **Table 1** Results of GAMs (beta regression family with logit link function) testing relationship
687 of local contribution to β-diversity ($LCBD_{D\%diff}$, $LCBD_{DS}$) with local species richness
688 (richness), and the relationship of species contribution to β-diversity ($SCBD_{ab}$, $SCBD_{pa}$) with
689 the number of sites occupied by a given species (occup) and the species' total abundance
690 (abund). edf = The estimated degrees of freedom accounting for smoothing function, Ref. df.
691 = Reference degrees of freedom, $\chi 2$ = Chi-square test statistic, adj. R^2 = The proportion of
692 variance explained by the model, Dev. expl. = The proportion of the null deviance explained
693 by the model, P = p-value
694

	smooth term	edf	Ref. df	χ^2	adj. R^2	Dev. expl. (%)	P
$LCBD_{D\%diff}$	s(richness)	1.000	1.000	2.09	0.032	5.6	0.148
$LCBD_{DS}$	s(richness)	4.034	4.915	68.06	0.644	64.6	<0.001
SCBDab	s(occup)+s(abund)	8.155	8.795	281.40			<0.001
		8.918	8.996	4171.20	0.996	99.1	<0.001
SCBDpa	s(occup)+s(abund)	8.437	8.891	4687.79			<0.001
		1.002	1.003	0.11	0.963	96.2	0.741

695



696

697 **Fig. 1** Summary of assumptions about the main processes in the four metacommunity

698 concepts (NT = neutral theory, PD = patch dynamics, ME = mass-effect, SS = species sorting)

699

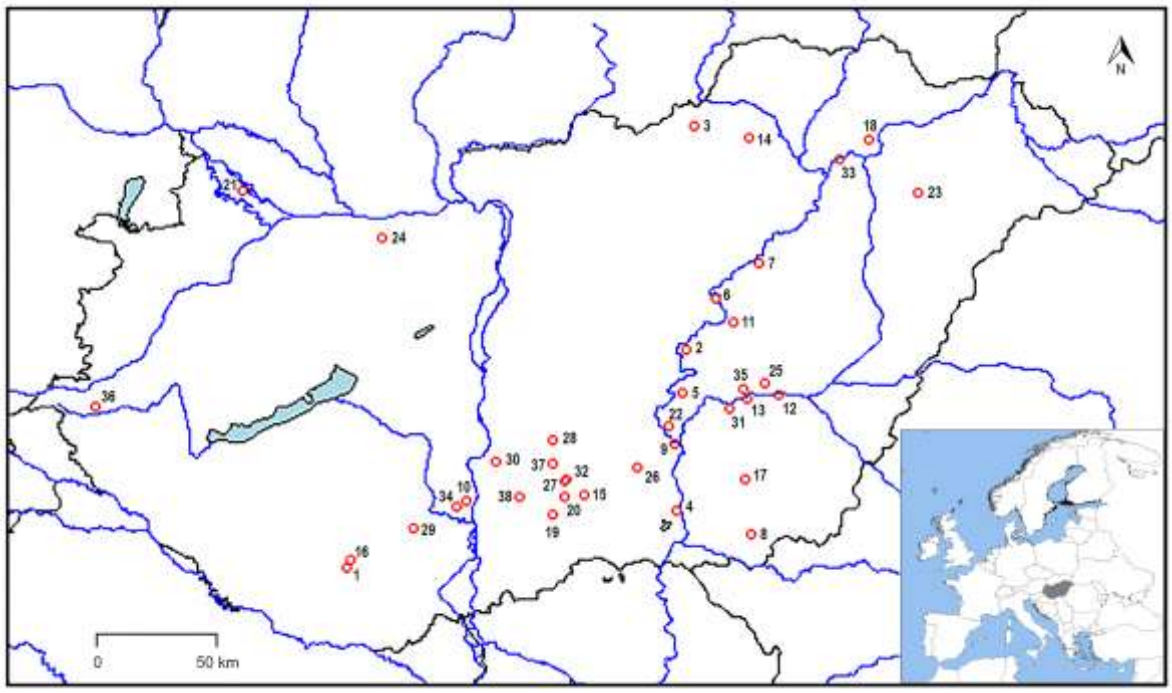


Fig. 2 Location and schematic map of Hungary and the sampling sites. Lake codes for the numbers are listed in Table S2

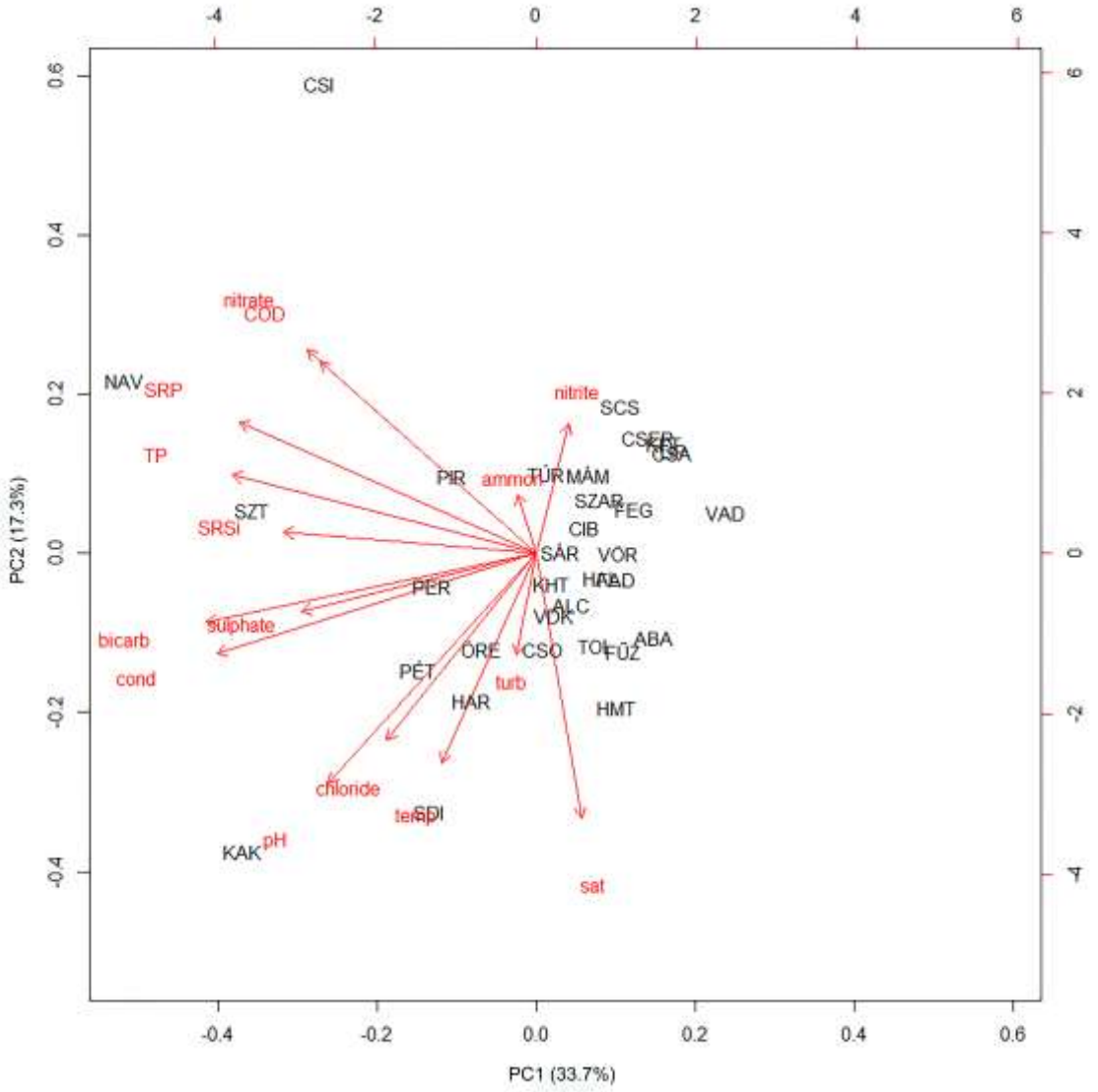
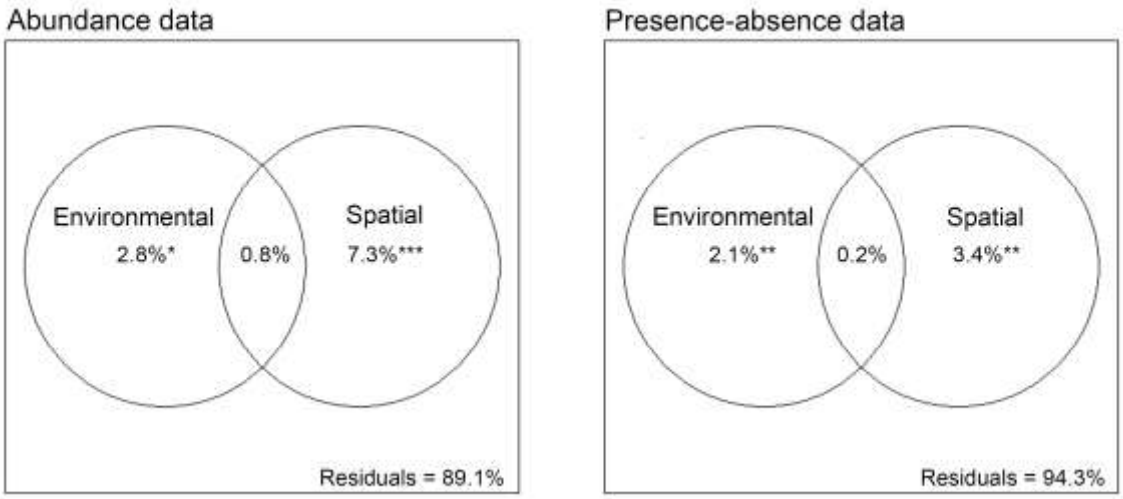


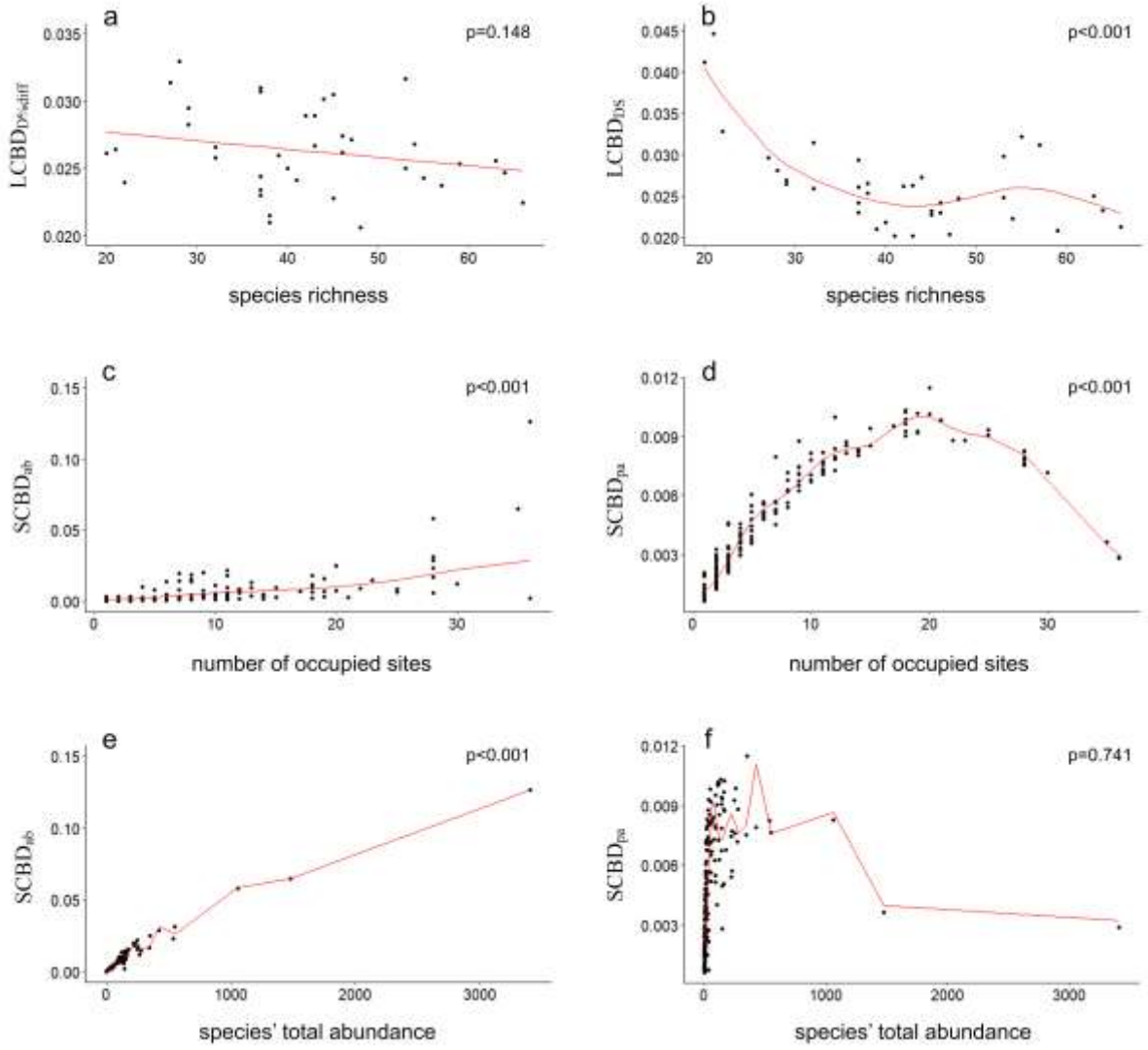
Fig. 3 PCA biplot representing sample codes and environmental variable vectors (cond = conductivity, sat = O_2 saturation, temp = water temperature, turb = turbidity, bicarb = HCO_3^-, chloride = Cl^-, sulphate = SO_4^{2-}, COD = chemical oxygen demand, nitrite = NO_2^-, nitrate = NO_3^-, ammon = NH_4^+, SRP = soluble reactive phosphorus, TP = total phosphorus, SRSi = soluble reactive silica). Proportion of variance explained by PC1 and PC2 axes are indicated



711

712 **Fig. 4** Results of variation partitioning conducted on Hellinger transformed abundance and
713 presence-absence data. Adjusted R^2 values are shown to indicate the relative importance of
714 environmental heterogeneity (Environmental) and spatial distance (Spatial) in the total
715 community variation. Unexplained variances are represented by the residuals. Significance of
716 testable fractions is shown as follows: *** = 0.001, ** = 0.01, * = 0.05. P values were
717 computed using ANOVA of RDA models

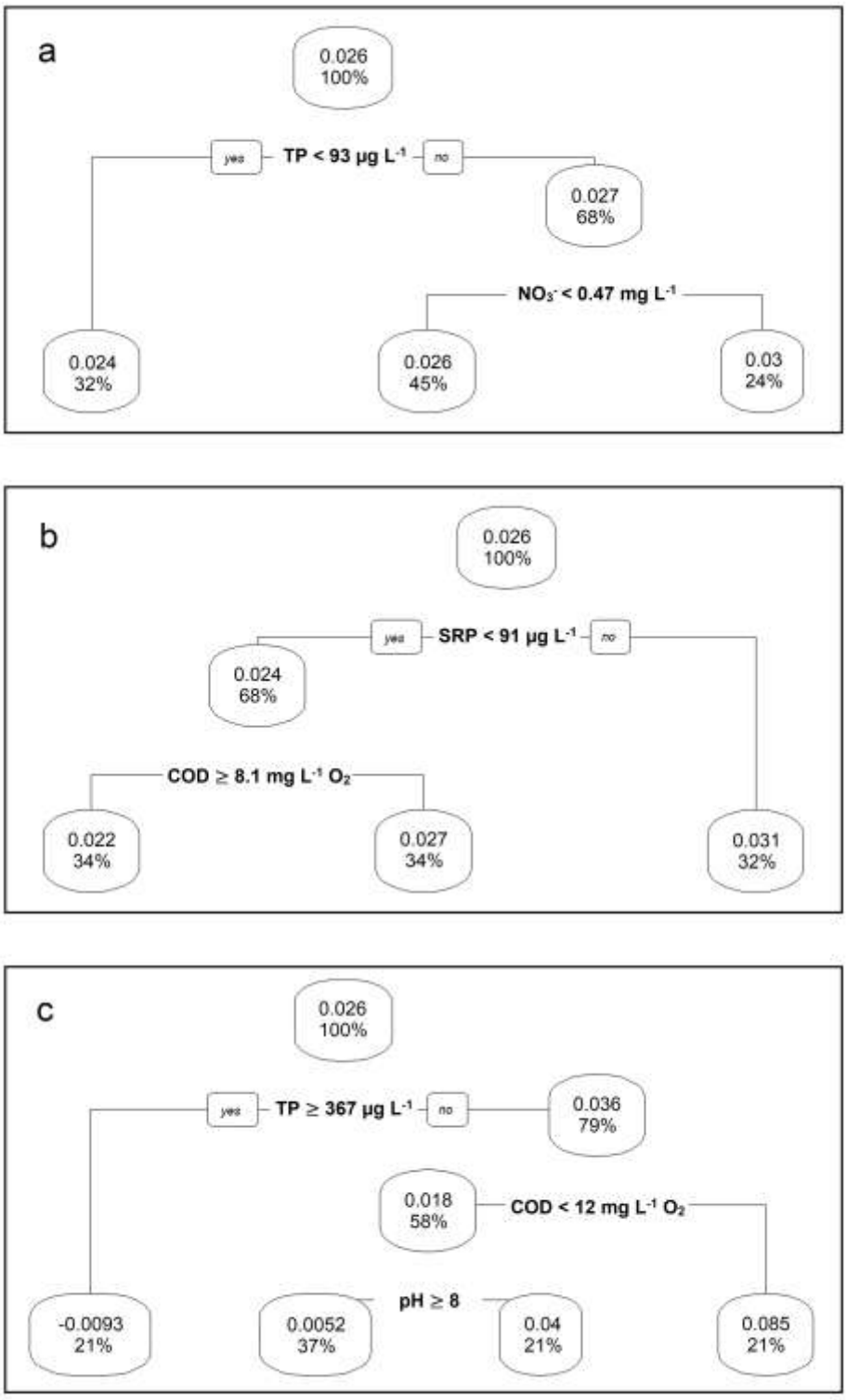
718



719

720 **Fig. 5** The relationship of local contribution to β-diversity ($LCBD_{D\%diff}$, $LCBD_{DS}$) with local
721 species richness and the relationship of species contribution to β-diversity ($SCBD_{ab}$, $SCBD_{pa}$)
722 with the number of occupied sampling sites and with the total abundance of a given species.
723 Solid lines show the fitted GAM using beta regression family with logit link function

724



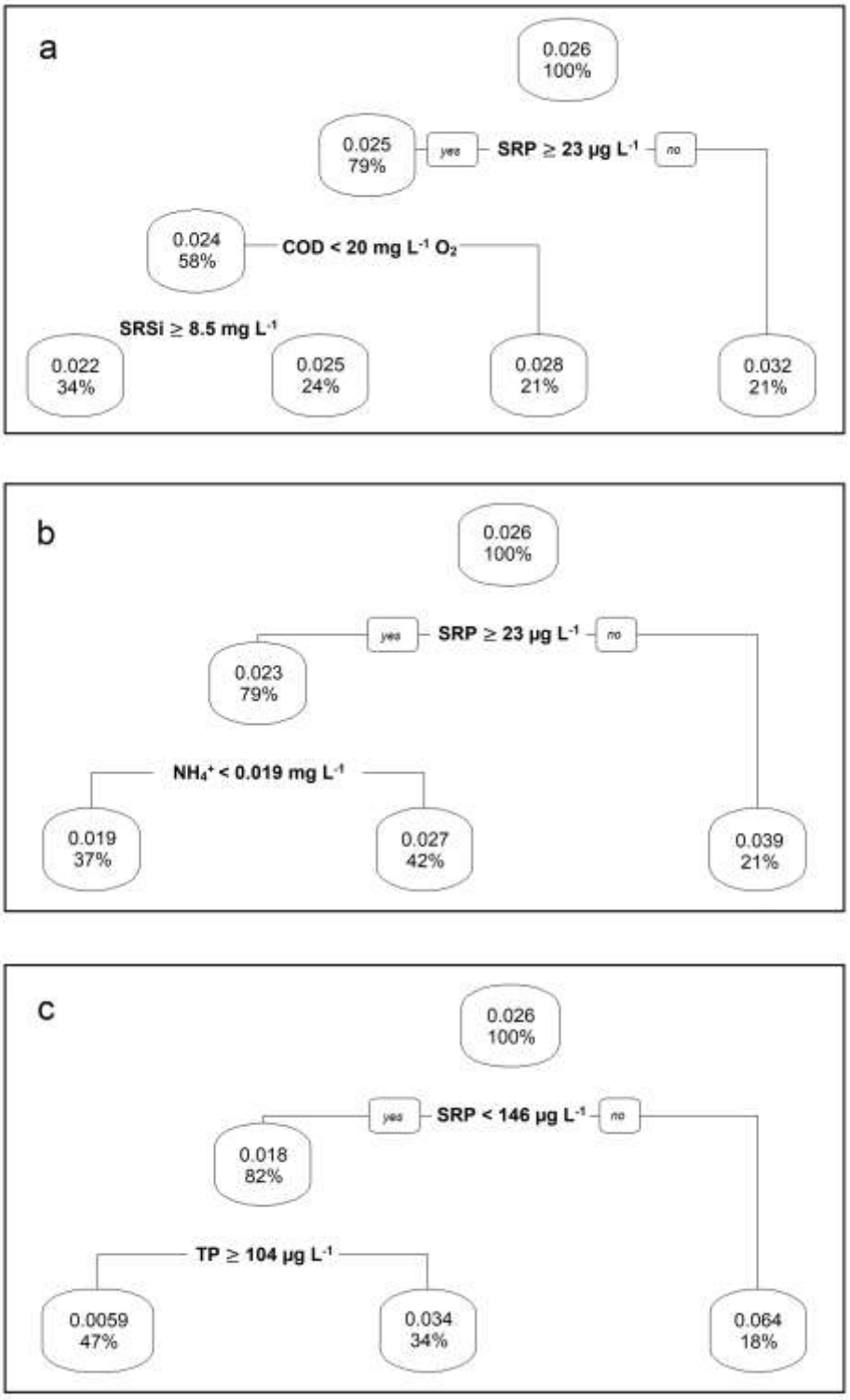
725

726 **Fig. 6** Regression trees for predicting a) $LCBD_{D\%diff}$, b) $LCBD_{ReplB\%diff}$ and c) $LCBD_{NesB\%diff}$

727 from the set of environmental parameters. Each node shows the predicted LCBD value (i.e.

728 the mean LCBD value) and the percentage of observations in the node

729



730

731 **Fig. 7** Regression trees for predicting a) $LCBD_{DS}$, b) $LCBD_{ReplBS}$ and c) $LCBD_{NesBS}$ from the
732 set of environmental parameters. Each node shows the predicted LCBD value (i.e. the mean
733 LCBD value) and the percentage of observations in the node